PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 5 :
C12N 15/12, C07K 15/06, 15/14

A1

(11) International Publication Number: WO 91/15583

(43) International Publication Date: 17 October 1991 (17.10.91)

(21) International Application Number: PCT/US91/02344

(22) International Filing Date: 4 April 1991 (04.04.91)

(30) Priority data:
504,800 5 April 1990 (05.04.90) US

(71) Applicant: THE AMERICAN NATIONAL RED CROSS [US/US]; 430 17th Street, N.W., Washington, DC 20006 (US).

(72) Inventors: HLA, Timothy, Tun ; 20424 Cabana Drive, Germantown, MD 20876 (US). MACIAG, Thomas ; 6050 Valerian Lane, Rockville, MD 20852 (US).

(74) Agents: GOLDSTEIN, Jorge, A. et al.; Sterne, Kessler, Goldstein & Fox, 1225 Connecticut Avenue, N.W., Suite 300, Washington, DC 20036 (US).

(81) Designated States: AT (European patent), AU, BB, BE (European patent), BF (OAPI patent), BG, BJ (OAPI patent), BR, CA, CF (OAPI patent), CG (OAPI patent), CH (European patent), CM (OAPI patent), DE (European patent), DK (European patent), ES (European patent), FI, FR (European patent), GA (OAPI patent), GB (European patent), GR (European patent), HU, IT (European patent), JP, KP, KR, LK, LU (European patent), MC, MG, ML (OAPI patent), MR (OAPI patent), MW, NL (European patent), NO, PL, RO, SD, SE (European patent), SN (OAPI patent), SU, TD (OAPI patent), TG (OAPI patent).

Published
With international search report.

(54) Title: A PROTEIN FAMILY RELATED TO IMMEDIATE-EARLY PROTEIN EXPRESSED BY HUMAN ENDOTHELIAL CELLS DURING DIFFERENTIATION

(57) Abstract

This invention provides a novel family of tissue specific genes and proteins that are related to a G-protein-coupled receptor gene and the receptor protein. The gene is an intermediate early gene that is expressed in differentiating endothelial cells. In particular, this invention provides a gene, edg-1, that is an immediate-early gene that encodes a G-protein-coupled receptor in endothelial cells. This invention also provides the G-protein-coupled receptor protein that is encoded by edg-1.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	**ES**	Spain	**MG**	Madagascar
AU	Australia	**FI**	Finland	**ML**	Mali
BB	Barbados	**FR**	France	**MN**	Mongolia
BE	Belgium	**GA**	Gabon	**MR**	Mauritania
BF	Burkina Faso	**GB**	United Kingdom	**MW**	Malawi
BG	Bulgaria	**GN**	Guinea	**NL**	Netherlands
BJ	Benin	**GR**	Greece	**NO**	Norway
BR	Brazil	**HU**	Hungary	**PL**	Poland
CA	Canada	**IT**	Italy	**RO**	Romania
CF	Central African Republic	**JP**	Japan	**SD**	Sudan
CG	Congo	**KP**	Democratic People's Republic of Korea	**SE**	Sweden
CH	Switzerland			**SN**	Senegal
CI	Côte d'Ivoire	**KR**	Republic of Korea	**SU**	Soviet Union
CM	Cameroon	**LI**	Liechtenstein	**TD**	Chad
CS	Czechoslovakia	**LK**	Sri Lanka	**TG**	Togo
DE	Germany	**LU**	Luxembourg	**US**	United States of America
DK	Denmark	**MC**	Monaco		

A PROTEIN FAMILY RELATED TO IMMEDIATE-EARLY PROTEIN EXPRESSED BY HUMAN ENDOTHELIAL CELLS DURING DIFFERENTIATION

BACKGROUND OF INVENTION

The endothelium is composed of a monolayer of quiescent cells, endothelial cells. Endothelial cells, which form the inner lining of blood vessels participate in a multiplicity of physiological functions, including the formation of a selective barrier for the translocation of blood constituents and macromolecules to underlying tissues and the maintenance of a non-thrombogenic interface between blood and tissue. Endothelial cells are also an important component in the development of new capillaries and blood vessels. Blood vessel development, which is called angiogenesis, occurs during developmental periods, such as during development of the vascular system, and as part of the pathophysiology of a variety of disease states, such as psoriasis, arthritis, chronic inflammatory conditions, diabetic retinopathy, and tumor development.

Angiogenesis, which involves the organized migration, proliferation, and differentiation of the endothelial cells, is initiated by the endothelial cell in response to angiogenic stimuli and can be separated into three distinct events: cell migration, cell proliferation and cell differentiation, whereby the cells organize into a tubular structure.

These events are mediated *in vitro*, and most likely *in vivo*, by mitogenic polypeptides. The migration of endothelial cells is induced by factors, including the heparin binding

growth factors and angiotropin. Proliferation is induced by the heparin binding growth factors (hereinafter HBGFs) and differentiation and cellular organization is induced by polypeptides, including interleukin-1 (hereinafter IL-1), tumor necrosis factor (hereinafter TNF), gamma-interferon, transforming growth factor alpha and beta (hereinafter TGF-α and TGF-β, respectively) and phorbol mistric acetate (hereinafter PMA).

The extracellular matrix (hereinafter ECM), which contains numerous components, also modulates endothelial cell differentiation. If endothelial cells are cultured in vitro on collagen gels in the presence of PMA organized networks of tubular structures form, and, if the cells are cultured in ECM conditioned medium the formation of tubular structures is accelerated.

The importance of the ECM components for mediation of endothelial cell differentiation is evidenced by the observations that antibodies that have been prepared against fibronectin and laminin inhibit formation of the differentiated phenotype, while proteolytic modification of fibronectin by plasmin leads to rapid modification of the endothelial cell phenotypic changes that are observed in vitro. In addition, competitive inhibitors of the laminin and fibronectin receptor binding domains also inhibit the ability of endothelial cells to complete the non-terminal differentiation program.

As discussed above, the polypeptide cytokines and PMA inhibit the HBGF-1-induced proliferation of endothelial cells and induce differentiation thereof. These factors induce a reversible phenotypic transition from a non-polar cobblestone monolayer into a polar elongated, fibroblast-like phenotype. The inhibition of HBGF-1-induced proliferation is mediated, at least in part, via down regulation of the HBGF-1 receptor.

It is also known that PMA activates protein kinase C, which a family of phospholipid- and calcium-activated protein kinases. This activation results in the transcription of an array of proto-oncogene transcription factors, including c-fos, c-myc and c-jun, proteases, protease inhibitors, including collagenase type I and plasminogen activator inhibitor, and adhesion molecules, including intercellular adhesion molecule I. Protein kinase C activation antagonizes growth factor activity by the rapid phosphorylation of the epidermal growth factor receptor. Phosphorylation decreases tyrosine kinase activity.

Upon induction of differentiation of endothelial cells in vitro by a cytokine or PMA, a set of immediate-early genes are rapidly induced via a pathway that does not require protein synthesis. Included among these immediate-early genes are transcriptional factors, cytokines, cytoskeletal proteins, nuclear hormone receptors and extracellular matrix receptors.

Cell surface receptors bind circulating signal polypeptides, such as growth factors and hormones, as the initiating step in the induction of numerous intracellular effector functions. Receptors are classified on the basis of the particular type of pathway that is induced. Included among these classes of receptors are those that bind growth factors and have intrinsic tyrosine kinase activity, such as the HBGF receptors and those that couple to effector proteins through guanine nucleotide binding regulatory proteins, hereinafter referred to as G-protein coupled receptors and G-proteins, respectively. The G-protein transmembrane signaling pathways consist of three proteins: receptors, G proteins and effectors.

G proteins, which are the intermediaries in transmembrane signaling pathways, are heterodimers and consist of α, β and gamma subunits. Among the members of a family of G proteins

the α subunits differ. Functions of G proteins are regulated by the cyclic association of GTP with the α subunit followed by hydrolysis of GTP to GDP and dissociation of GDP.

G-protein coupled receptors are a diverse class of receptors that mediate signal transduction by binding to G-proteins. Signal transduction is initiated via ligand binding to the cell membrane receptor, which stimulates binding of the receptor to the G-protein. The receptor-G-protein interaction releases GDP, which is specifically bound to the G-protein, and permits the binding of GTP, which activates the G-protein. Activated G-protein dissociates from the receptor and activates the effector protein, which regulates the intracellular levels of specific second messengers. Examples of such effector proteins include adenylyl cyclase, guanylyl cyclase, phospholipase C, and others.

G-protein-coupled receptors, which are glycoproteins, are known to share certain structural similarities and homologies (see, e.g., Gilman, A.G., Ann. Rev. Biochem.56: 615-649 (1987), Strader, C.D. et al. The FASEB Journal 3: 1825-1832 (1989), Kobilka, B.K., et al. Nature 329: 75-79 (1985) and Young et al. Cell 45: 711-719 (1986)). Among the G-protein-coupled receptors that have been identified and cloned are the substance K receptor, the angiotensin receptor, the α- and β-adrenergic receptors and the serotinin receptors. G-protein-coupled receptors share a conserved structural motif. The general and common structural features of the G-protein-coupled receptors are the existence of seven hydrophobic stretches of about 20-25 amino acids each surrounded by eight hydrophilic regions of variable length. It has been postulated that each of the seven hydrophobic regions forms a transmembrane α helix and the intervening hydrophilic regions form alternately intracellularly and extracellularly

exposed loops. The third cytosolic loop between transmembrane domains five and six is the intracellular domain responsible for the interaction with G-protein.

G-protein-coupled receptors are known to be inducible. This inducibility was originally described in lower eukaryotes. For example, the cAMP receptor of the cellular slime mold, Dictyostelium, is induced during differentiation (Klein et al., Science 241: 1467-1472 (1988). During the Dictyostelium discoideum differentiation pathway, cAMP, induces high level expression of its G-protein-coupled receptor. This receptor transduces the signal to induce the expression of the other genes involved in chemotaxis, which permits multicellular aggregates to align, organize and form stalks (see, Firtel, R.A., et al. Cell 58: 235-239 (1989) and Devreotes, P., Science 245: 1054-1058 (1989)). Human endothelial cells utilize a series of morphological correlates during its differentiation pathway, discussed supra., in which individual cells migrate, align and organize to form multicellular capillary-like structures.

SUMMARY OF THE INVENTION

It is one object of this invention to provide a novel G-protein-coupled receptor that is the product of an immediate early gene that is expressed in endothelial cells during the early stage of differentiation.

It is another object of this invention to provide a family of proteins that are expressed in a tissue-specific manner and that are related to the novel G-protein-coupled receptor that is the product of an immediate early gene that is expressed in endothelial cells during the early stage of differentiation.

It is another object of this invention to provide DNA molecules that encode each member of the family of proteins that are expressed in a tissue-specific manner and that are related to the novel G-protein-coupled receptor that is the product of an immediate early gene that is expressed in endothelial cells during the early stage of differentiation.

It is another object of this invention to provide DNA molecules that encode the novel G-protein-coupled receptor that is the product of an immediate early gene that is expressed in endothelial cells during the early stage of differentiation.

In accordance with this invention there is provided a DNA molecule that encodes edg-1 gene product, which is the product of an immediate-early gene that is expressed in the early stage of differentiation of endothelial cells in response to PMA or IL-1.

This invention provides a gene and protein, which is the first immediate-early gene that encodes a G-protein-coupled receptor.

Unless defined otherwise, all technical and scientific terms used herein have the same meaning as is commonly understood by one of ordinary skill in the art to which this invention belongs. Although methods and materials similar or equivalent to those described herein can be used in the practice of testing of the present invention, the preferred methods and materials are now described. All publications mentioned hereunder are incorporated by reference.

BRIEF DESCRIPTION OF THE FIGURES

Figure 1. The identification of edg-1, an Immediate early gene induced by PMA in HUVEC (human umbilical vein endothelial cells).

Confluent cultures of HUVEC were treated with 20 ng/ml of PMA for the indicated times. The cells were then lysed, RNA purified, and total RNA (10 μg) analyzed by Northern blot analysis. The cDNA probes that were used were edg-1 (A) and glyceraldehyde-3-phosphate (GAPDH) (B) cDNA.

Figure 2. Confluent cultures of HUVEC were treated with the indicated reagents for 4 hour and the RNA was isolated. Total RNA (10 μg) was fractionated by 1% agarose-formaldehyde gel electrophoresis, blotted onto a zeta-probe membrane and hybridized with [^{32}P]-labeled edg-1 (A) or a GAPDH (B) cDNA probes. The following reagents were used: PMA (20 ng/ml), chx (5 μg/ml), Actinomycin D (Act D) (2 μg/ml). Each reagent was used either alone or in combination.

Figure 3. Confluent cultures of HUVEC were pre-treated with 20 ng/ml PMA for 4 hour. Either Act D (2 μg/) alone or with chx (5 μg/ml) was added to the cultures, at a time designated 0. At the indicated time points, cultures were harvested and Northern blot analysis was performed on total RNA as described above using the edg-1 (A) and GAPDH (B) cDNA probes.

Figure 4. HUVEC were either untreated or treated with 20 ng/ml PMA for 2 hour after which nuclei were prepared. Run-off transcripts were obtained by labelling 10^7 nuclei *in vitro* with [^{32}P]-UTP. RNA was purified and hybridized to immobilized plasmid DNA encoding edg-1 (10 μg/slot), human fibronectin (fn) (2 μg/slot) and pBluescript (pBS) (10 μg/slot).

Figure 5. Nucleotide and Deduced Amino Acid Sequence of Human edg-1.

The nucleotide (1-2774) and deduced amino acid sequence (1-380) is shown for human edg-1 cDNA. The deduced transmembrane domains are underline and potential N-linked glycosylation sites are shown with ann asterisk. Possible serine and threonine phosphorylation sites are shown with closed circles. The basic amino acid-rich intracellular domain, which is located between transmembrane domains five and six is highlighted with open circles. THe Kozak consensus translation initiation sequence (5') and polyadenylation sites (3') are shown with double lines underneath their respective sequences. The Genbank accession number for this nucleotide sequence is M31210.

Figure 6. The amino acid sequence of the putative edg-1 translation product was aligned with Substance K receptor (SKR), Substance P receptor (SPR), β_2-adrenergic receptor (B2AR), Serotonin receptor 1c (5HTC), α_2-adrenergic receptor (A2A), Serotonin receptor 1a (5HT1a), Rhodopsin (OSPD) and angiotensin receptor (MAS). Highly homologous regions are boxed and indicated on the linear schematic.

Figure 7. A structural model for the putative edg-1 translation product is shown. This model is analogous to other G-protein-coupled receptors. The potential N-linked glycosylation sites are indicated with an inverted "Y". Potential phosphorylation sites at serine and threonine residues are shown with dark circles. The third cytosolic intracellular domain, which is between transmembrane domains 5 and 6 contains a highly basic region (11/35 residues) is also indicated.

Figure 8. Hydrophobicity Profile of edg-1 Translation Product. The deduced amino acid sequence of edg-1 was analyzed for hydrophobic regions and the amino acid sequence (residues) plotted against the hydrophobicity index. The putative transmembrane (TM) domains are indicated.

Figure 9. Expression of edg-1 transcript in human cells. Total RNA (5 μg) from human saphenous vein smooth muscle cells (S), foreskin fibroblasts (F), HeLa cells (H), epidermoid carcinoma (A431) cells (A), melanocytes (M), brain tissue (B) and endothelial cells (E) were reverse transcribed into cDNA and amplified with edg-1 specific oligonucleotide primers that span the carboxy-terminal tail domain (A) and the third cytosolic loop (B). Amplified DNA was separated by agarose gel electrophoresis and visualized by ethidium bromide staining. Molecular weight markers (indicated by arrows) are from top to bottom: 1.6 Kb, 1.0 Kb, 0.5 Kb, 0.4 Kb, 0.3 Kb, 0.2 Kb and 0.15 Kb.

It can be seen in (A) that transcript of the expected size, about 600 bp,, which was amplified using oligonucleotide primers specific for the C-terminal domain, was present in RNA from all the cultured cell lines and human brain. In contrast, when the transcript was amplified using an a pair oligonucleotides that span the third intracellular loop, cell or tissue specific bands were observed.

DESCRIPTION OF THE PREFERRED EMBODIMENTS

In the invention described herein a novel gene, edg, and the protein encoded thereby has been identified. In addition, this invention provides a family of proteins that are structurally and functionally related to this protein as well as DNA molecules, but that are tissue or cell type specific are provided.

As used herein, the edg-G-protein-coupled receptor family is a family of related proteins that share substantial homology and structure and that contain common constant regions or domains but differ in at least one variable region or domain that includes the third cytosolic loop. See, e.g., Figures 6, 7, and 9. The particular variable region and, thus, each family member, is expressed in a tissue-specific manner.

As used herein, expression of a transcript in a tissue-specific manner includes expression of transcripts that are expressed in only certain tissues or cell types. Such tissue-specific expression can be effected through a variety of mechanisms, including the expression of different genes in each tissue or cell type, through alternative splicing of the same gene in each tissue or cell type, or through recombination of germ line DNA in during development or differentiation of each cell type.

As used herein, the edg-1-G-coupled protein receptor transcript is the intermediate early transcript that is expressed in the early stage of differentiation in endothelial cells that can be induced or stimulated with PMA and interleukin-1 (IL-1) but not with TGF-β, HBGF-1, or α-thrombin. The edg-1 G-coupled protein receptor transcript encodes the edg-1 G-coupled protein receptor.

As used herein, the edg-1-G-coupled protein receptor transcript family is a family of transcripts that are expressed in a tissue-specific manner and encode members of the family of related proteins that share substantial homology and structure and that contain common constant regions or domains but differ in at least one variable region that includes the third cytosolic loop.

As used herein, DNA encoding a protein includes any DNA molecule that encodes a protein that has substantially the same amino acid sequence. Each of such proteins may, however, differ at sites that are not essential to protein function and includes proteins isolated from different individuals in the same species, proteins isolated from different species that share substantially the same biological activities, and proteins isolated from different cultured cell lines.

As used herein, the edg-1 transcript refers to the 2.8 Kb (about 3 Kb) transcript that encodes the receptor protein. This term is herein used interchangeably with the edg transcript, edg mRNA. The edg-1 transcript also refers to this transcript, but also refers to the 1-Kb clone that was isolated from the differential screen, which contained a poly A tract at 3' end, a unique nucleotide sequence and hybridized to the about 3.0 Kb PMA inducible mRNA species, the edg-1 transcript.

Because PMA inhibits endothelial cell proliferation and induces differentiation, the identification and isolation of immediate-early genes yields insight into the molecular mechanisms involved in the regulation of endothelial cell differentiation.

Immediate-early genes that are expressed in endothelial cells may be isolated from any source of endothelial RNA. In one embodiment of this invention, human umbilical vein endothelial cells (hereinafter HUVEC) are used. The HUVEC are either untreated and treated with PMA, IL-2 or any other signal that induces these genes.

The desired immediate-early genes can be identified by any means in which the transcripts comparing the transcripts in cells that are stimulated with PMA, IL-2 or other inducer with the transcripts that are present in untreated cells. Those that are present only in the treated cells are, thus, immediate-early genes. In addition, any member of the G-protein-coupled receptor family of this invention can be identified by screening an appropriate library with an appropriate probe derived from the edg-1 clone. For example, an appropriate probe would be one derived from the 3' end of the clone. Any methods known to those of skill in the art to accomplish this may be used.

In endothelial cells the immediate-early gene of this invention is the edg-1 encoding gene. It is induced by IL-1, LPS or PMA, but not by HBGF-1, TGF-β, or α-thrombin. The edg-1 clone provided herein encodes a protein that shares many structural and sequence similarities with known G-protein-coupled receptors, including the β-adrenergic, substance K, substance P, rhodopsin, serotonin (5-HT), tachykinin receptors and the cAMP receptor of Dictyostelium.

The N-linked glycoslyation site at Asn_{30} is also found in the Substance K and angiotensin receptors. The two N-linked glycosylation sites are found within the amino-terminal domain of all G-protein-coupled receptors. The region in proximity to the second and third hydrophobic domains is highly conserved among all such receptors, including that encoded by edg-1. In the β_2-adrenergic receptor Asp_{130} is known to be absolutely necessary for G-protein; in the edg-1-encoded protein the Asp/Glu-Arg is conserved.

Although the overall sequence similarity between the edg-1 G-protein-coupled receptor of this invention and other such receptor is quite divergent, there is a significant degree of sequence similarity within the carboxy-terminal

half, particularly within transmembrane domain seven. It is most similar to those receptors that recognize peptides as receptor ligands.

The intracellular hydrophilic loop regions contain four potential phosphorylation sites at residues Thr_{72}, Ser_{231}, Thr_{235} and at Ser_{351}. This feature is common to many G-protein-coupled receptors. Phosphorylation at the Ser and Thr residues within the intracellular domains has been implicated in the phenomenon of receptor desensitization.

The hydrophilic region between transmembrane domains five and six is the region that is absolutely necessary for G-protein coupling and it is highly divergent among members of the G-protein-coupled receptor proteins. In the G-protein-coupled receptor that is encoded by edg-1, this region is highly basic. The family of edg-1 related tissue-specific proteins provided in this invention differ in this region and, thus, most likely differ in their respective binding or coupling interactions with the G-protein or protein ligands.

The ligand that binds to each of the members of the family of G-protein-coupled receptor proteins of this invention can be identified by methods that are known to those of skill in the art. For example, xenopus oocytes can be transfected with DNA that encodes the particular protein. The protein will be expressed on the cell surface of the oocytes. Since these oocytes are sensitive to calcium exchange across the cell membrane, binding of the appropriate ligand causes calcium exchange across membrane. Labeled calcium can be used and the ligand that causes labeled calcium exchange can be identified. Among the candidates for the ligand that binds to the edg-1-G-protein coupled receptor are ATP, AMP, adenosine, leukotrienes, prostenoids, histamine, bombasin, thrombin, azopressin, bradykinin, endothelin, serotensin, substance P and neuropeptide.

The following examples are included for illustrative purposes only and are not intended to limit the scope of the invention.

EXAMPLE 1

Materials and Cell Culture

Recombinant human interleukin aα (IL-1α), which was the gift of Dr. Peter Lomedico, Hoffman La Roche, Nutley, NJ. Recombinant human HBGF-1α was obtained from Anthony Jackson, American Red Cross, Rockville, MD. Porcine TGF-β was purchased from R & D Systems.

Primary cultures of human umbilical vein endothelial cells (HUVEC) were obtained from Dr. MIchael Gimbrone, Harvard Medical School, Boston, MA ,and were grown on fibronectin-coated plates in Medium 199 supplemented with 10% (v/v) fetal bovine serum, 1x antiobiotic and antimycotic mixture (GIBCO, Grand Island, NY), 150 µg/ml crude endothelial cell growth factor (Maciag et al. , 1981) and 5 U/ml heparin (Sigma) as described in Maciag et al. ((1981) J. Biol. Chem. 91, 420-426). Cells were subcultured at a 1:5 split ratio and cultures between passages of 4 and 12 were used. At confluence, cells were maintained in medium without the growth factor and heparin for two days to achieve quiescence.

RNA Preparation and cDNA Library Construction

Total RNA was obtained from cells that either untreated or treated with 20 ng/ml PMA (Sigma) and 5 µg/ml of cycloheximide (hereinafter chx) (Sigma) for 4 hours. The cells were rinsed with phosphate-buffered saline, lysed in 4M guanidinium isothiocyanate and total RNA purified as described in Winkles, J., et al. ((1987) Proc. Natl. Acad. Sci. USA 84,

7124-7128). Poly A^+ RNA (10 μg) from HUVEC exposed to PMA and chx was converted to double-stranded cDNA and cloned into the Eco R1 site of lambda gt10, using the cDNA synthesis kit from Bethesda Research Labs (Gaithersburg, MD) and the cDNA cloning kit from Amersham (Chicago, IL). The library contained > 10^6 independent clones, with an average insert size of approximately 1 Kb.

Northern Blot Analysis.

Total RNA (10 μg) was electrophoresed on a 1% agarose gel containing 2.2 M formaldehyde, capillary-blotted onto Zeta-probe membrane (Biorad) and UV cross-linked (Maniatis et al. (1982) In Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY). The cDNA insert fragment for edg-1 (2.8 Kb) or human GAPDH (1 Kb) was labeled to high specific activity (>108 cpm/μg) using a random primer labeling kit (BRL) and was used to hybridize filters in Church-Gilbert buffer (0.5 M sodium phosphate pH 7.2, containing 7% SDS and 1% bovine serum albumin, 1mM EDTA and 20% formamide at 65° C for 16-20 hrs. Filters were washed twice for 15 min at high-stringency (0.1xSSC, 65° C).

Differential Screening of cDNA Library

The differential screen was performed by plating 2 x 10^4 pfu of the library onto bacteriological plagues (15 cm diameter) containing LB agar. The phage were allowed to grow at 37° C until plaques were approximately 0.5 mm in diameter. Phage DNA was adsorbed onto Gene-screen plus nylon filters (Dupont, DE), in duplicate, denatured, neutralized, and UV cross-linked.

The probe for differential screening was prepared by reverse transcription of 1 μg of poly A^+ RNA from control and PMA/chx-treated HUVEC. The reaction conditions were as follows: 50 mM Tris HCl, pH 8.3, 75 mM KCl, 20 mM dithiothreitol, 3 mM $MgCl_2$, 500 μCi[^{32}P]-α-dCTP, 20 μM dCTP, 200 μM each of dATP, dCTP, and dTTP, 0.5 μg/ml of oligo dT_{12-18} and 400 units of MMLV-reverse transcriptase (Bethesda Research Labs, Gaithersburg, MD).

After incubation at 37° C for 60 minutes, RNA was hydrolyzed by treatment with 100 μl 0.6M NaOH and 20 mM EDTA for 30 minutes at 65° C. The cDNA was purified on Sephadex G-50 columns and ethanol-precipitated. Duplicate filters were incubated with 10 cpm/ml of cDNA for 48 hours at 65° C in hybridization buffer containing 2% SDS, 1 M NaCl and 10% dextran sulfate. The filters were washed twice for 30 min at 65° C with 2xSSC containing, 1% SDS followed by two additional washes for 30 min at 65° C with O.1xSSC containing 1% SDS.

The filters were autoradiographed and duplicates were superimposed on each other to isolate PMA/chx-induced signals. Differential signals were plaque-purified by repeating the screening process. Insert cDNA was prepared and used for either Northern blot analysis or subcloning into plasmid vectors.

Of the twelve positive signals obtained from $>10^5$ pfu of the library three were found to be consistently positive. Two of the clones had inserts had sequences identical to the sequence of DNA that encodes human collagenase Type 1. The third clone, herein called edg-1 (1-Kb) contained a poly A tract at 3' end, a unique nucleotide sequence and hybridized to a 3.0 Kb PMA inducible mRNA species.

This 1 kb insert was used to rescreen two additional cDNA libraries-lambda gt10 and cDM8. The largest clone was 2.8 kb. Further investigation and analysis was conducted using this clone, which is expressed at high levels (0.05%) in the HUVEC.

EXAMPLE 2

The kinetics of edg RNA induction by PMA was studied by Northern blot analysis of HUVEC that were exposed to PMA for 0.5, 1, 2, and 4 hours (Figure 1 (A)).

In order to determine the characteristics of the rapid edg-1 induction, Northern blot analysis was performed with HUVEC that had been treated for 4 hours with PMA and chx, alone or in combination (Figure 2). As can be seen in Figure 2, the 3.0 KB mRNA edg transcript was induced independently by PMA and chx, but was superinduced in the presence of both.

EXAMPLE 3

Chx was shown to exert the superinduction effect by stabilizing the edg-1 transcript (Figure 3). HUVEC were stimulated for 4 hour with PMA and subsequently incubated with actinomycin D, in inhibitor of transcription both in the presence and absence of chx. As shown in Figure 3 steady-state levels of the edg-1 mRNA declined to undetectable levels two hours after the addition of actinomycin D; whereas, chx prevented this decline.

EXAMPLE 4

In order to ascertain at what level PMA induces edg-1 mRNA, edg 1 induction in the presence of actinomycin D was investigated. As shown in Figure 2, actinomycin D repressed the inductive effect of PMA, which suggests that PMA induces the transcription of the edg-1 gene.

EXAMPLE 5

Nuclear Run-On Transcription.

Nuclei (10^7) were prepared from quiescent HUVEC untreated or treated with 20 ng/ml PMA for 2 hr. In vitro labeled, run-off transcripts were prepared by incubating the nuclei with 250 μCi of[α -^{32}P]-UTP (.6000 CI/mmol, Amersham), 10mM ATP, CTP, GTP, in the reaction buffer containing 20mM Tris-HCl, pH 7.9, 140mM KCl, 10mM $MgCl_2$ and 1mM dithiothreitol as described (Nevins, J., (1987) Meth. Enzymol. 152, 234-240).

The labeled RNA was purified (Winkles, J., supra.) and hybridized to nylon filters containing either 10 μg of denatured plasmid edg-1 cDNA, 2 μg of human fibronectin or 10 μg of pBluescript (Stratagene). The hybridization and washing conditions were identical to those described for the differential hybridization.

Nuclei were prepared from untreated HUVEC or from HUVEC treated with PMA for 2 hours. Labeled run-on transcripts were obtained and hybridized to immobilized plasmid DNA containing the edg-1 insert and to a control plasmid containing fibronectin-encoding DNA or to a Bluescript plasmid (Figure 4). Edg-1 transcription was significantly induced in nuclei from the PMA treated HUVEC.

EXAMPLE 6

DNA Sequence Analysis.

The structure of the edg-1 gene and gene product was elucidated by DNA sequencing of the 2.8 Kb cDNA clone.

Plasmid DNA for edg-1 (2.8Kb) was obtained by screening a cDNA library from HUVEC constructed in the vector, cDM8, which was a gift of Brian See, Harvard Medical School) with the (1.6Kb) insert obtained from the cDNA library in lambda

gt10, discussed in Example 1. Double-stranded sequence analysis was performed using the sequenase-2 enzyme (USBC), following the manufacturer's instructions. Successive primers were synthesized and used to sequence both strands of the cDNA clone. The DNA sequence was analyzed by the Intelligenetics Sequence Analysis program.

As shown in Figure 5, the complete nucleotide sequence of the edg-1 cDNA clone is 2774 bp long and, at nucleotide 251 from the 5'end, contains a consensus translation initiation sequence, which is followed by an open-reading frame (ORF) that encodes 380 amino acids. The ORF is followed by a 3', A/T-rich, 1.3 Kb untranslated region followed by a poly A tail. A/T rich sequence motifs in 3' untranslated regions have been implicated in conferring rapid RNA degradation of intermediate-early mRNAs. There are two consensus polyadenylation sites (AATAAA) at nucleotides 2590 and 2737, respectively. The edg-1 clone also contains about 250 bp of 5'untranslated region.

The deduced amino acid sequence contains a non-hydrophobic amino-terminal stretch of 46 amino acids, which contain two potential N-linked glycosylation sites at residues 29 and 35. This stretch is followed by seven alternating stretches of hydrophobic regions, each about 20 amino acid residues long. There are 8 hydrophilic regions. Each of the hydrophobic regions is flanked by hydrophilic regions of 7 to 19 amino acids, except for the region between the fifth and sixth transmembrane domain, which is 35 residues long and is rich in basic and dibasic residues. The last transmembrane domain is followed by a long, 66 amino acid, stretch of hydrophilic residues that include an abundance of serine and threonine residues.

EXAMPLE 7

Reverse Transcriptase-Polymerase Chain Reaction Analysis

RNA from HUVEC was purified as described in Example 1. RNA from human saphenous vein smooth muscle cells, human foreskin fibroblasts, human epidermoid carcinoma cells (A431), human cervical carcinoma cells (HeLa), human melanocytes and total brain were the generous gift of Dr. Jeffrey Winkles of the American National Red Cross.

Total RNA (5 μg) from all the cultured cells and poly A^+RNA (1 μg) from human brain (Clontech) was converted to cDNA by treatment with 200 units of MMLV reverse transcriptase (Bethesda Research Labs, MD) in 50 mM Tris-HCl, pH, 8.0, 1 mM dithiothreitol, 15 mM NaCl, 3 mM $MgCl_2$, 1 unit RNAsin (Promega), o.2 μg of random hexamer primers, 0.8 mM dNTPs and incubated for 1 hour at 37° C. The reaction was terminated by heating at 95° C for 10 minutes and diluted to 1 ml with distilled water.

Enzymatic amplification was done on a 10 μl aliquot of the cDNA mix. PCR was performed in 50 mM Tris-HCl, pH, 8.0, 1.5 mM $MgCl_2$, 10 mM KCl, 0.2 mM dNTPs, 0.5 μg each of primers for edg-1 and 2.5 units of Taq DNA polymerase (Cetus, CA) (see, Saiki et al. (1988) Science 239, 487-491). The reaction mixture was heated at 94° C for 1 minute, annealed at 55° C for 2 minutes, and extended at 72° C for 3 minutes for 30 repetitive cycles. The primers used were as follows:

(1) 5'-TG TAC TGC AGA ATC TAC T-3' (sense) and 5'-T GCA GCC CAC ATC CAG CAG CA-3' (antisense) to amplify from nucleotide no. 909 to 1094, which spans the third cytosolic domain; and

(2) 5' AAG ACC TGT CAC ATC CTC TTC-3' (sense) and 5' ATG AAC CCT TTA GGA GCT TGA CAA-3'(antisense) to amplify from nucleotide no. 1100 to 1702, which spans the seventh transmembrane domain, the cytosolic tail and part of the 3'untranslated region.

When RNA from the various cultured human cell lines and from human brain was reverse transcribed and the cDNAs amplified using the oligonucleotides that are specific for the C-terminal domain (amino acids 266 to the termination codon and 309 bp of the 3' untranslated region, nucleotides 1100 to 1702, see, e.g., Figures 5-7 and 9) an amplified product is the expected size, 600 bp., is observed (see Fig. 9 (A)) in RNA from all cell types and human brain. The intensity of the signal was most prominent in endothelial cells, but was present to a lesser extent in smooth muscle cells, fibroblasts, epidermoid cells, melanocytes, and brain tissue.

When the cDNAs were amplified with a pair of oligonucleotides that span the third intracellular loop (amino acids 220-282, nucleotides 909-1094), cell-specific bands were amplified (Figure 9 (B)). In smooth muscle cells, a major band at 0.7 Kb and minor bands at 0.9, 0.3, and 0.19 Kb were observed. In HeLa cells a very prominent band was observed at 0.3 Kb. The expected 0.19Kb amplification product was observed only in endothelial cells.

This result indicates that cDNAs derived from mRNAs that are related to, but not identical with, the edg-1 transcript are present in different cell types and tissues. Because the third cytosolic loop has been identified in other G-protein-coupled receptors as the region that binds to the G-protein, the tissue specific transcripts differ in the region that encodes the portion of the receptor that couples with the G-protein and thereby modulates the cellular response of the particular cell type to the specific signal.

Since modifications will be apparent to those of skill in the art, it is intended that this invention be limited only by the scope of the appended claims.

We claim:

1. A purified DNA molecule that encodes a protein having the sequence of amino acids set forth in Figure 5.

2. The purified DNA molecule having the sequence of nucleotide bases set forth in Figure 5.

3. A purified protein that has substantially the same amino acid sequence as the sequence of amino acids set forth in Figure 5.

4. A purified DNA molecule that encodes the protein of claim 3.

5. A protein that includes regions that are substantially homologous with all or a portion of the protein of Figure 5, wherein said portion consists of the amino acids that comprise the transmembrane domains of the protein of FIgure 5.

6. A protein selected from the group consisting of the edg-1-G-coupled-protein receptor family of proteins.

7. The protein of claim 6, that is expressed in a cell or tissue selected from the group consisting of smooth muscle cells, fibroblasts, cultured immortal human cell lines, epidermoid carcinoma cells, melanocytes, brain tissue and differentiating endothelial cells.

8. An isolated DNA molecule that encodes the protein of claim 7.

FIG. 1

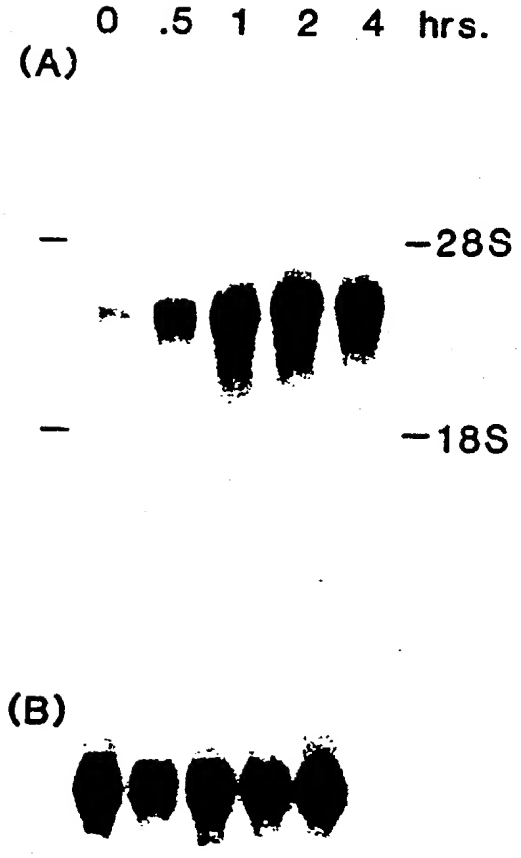


FIG. 2

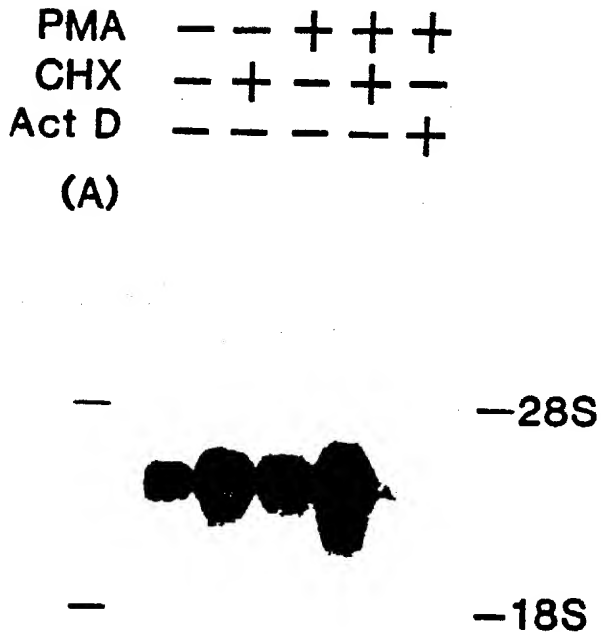


FIG. 3

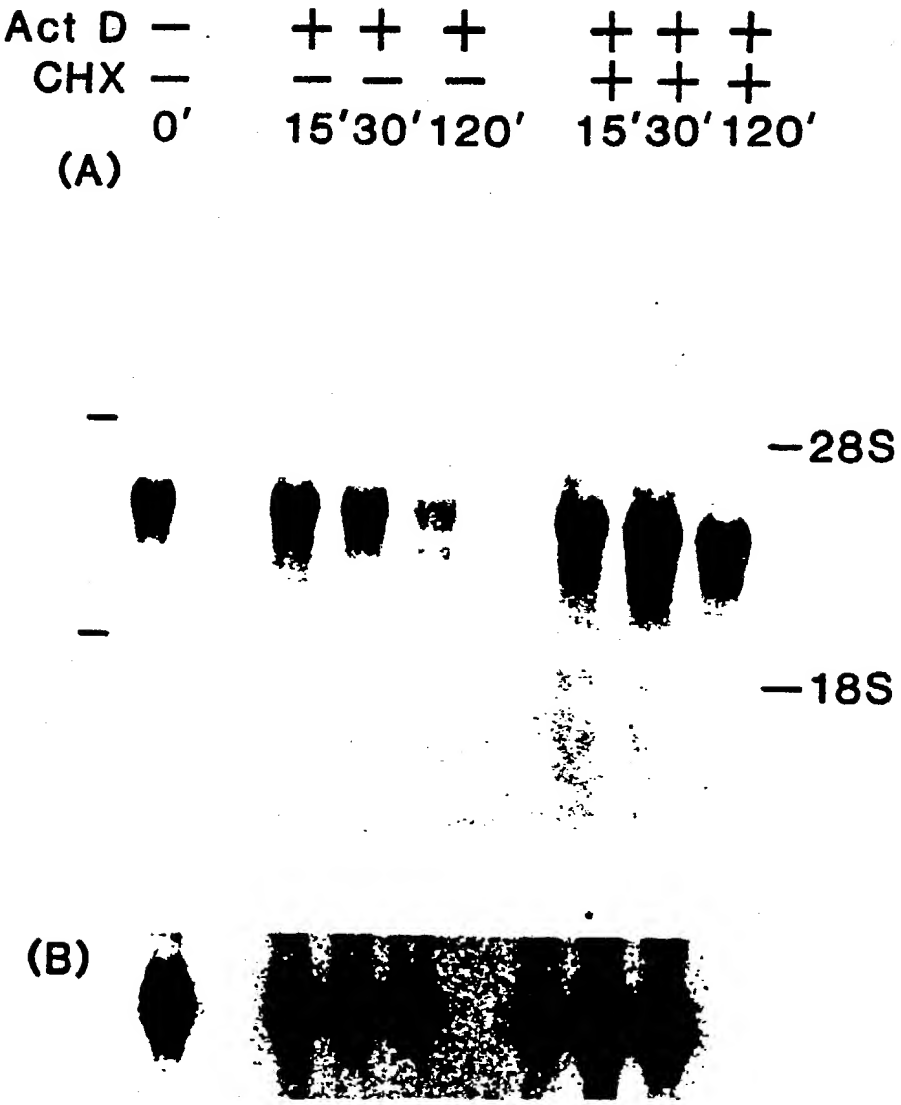


FIG. 4

```
        10          20          30          40          50          60          70
TCTAAAGGTC GGGGGCAGCA GCAAGATGCG AAGCGAGCCG TACAGATCCC GGGCTCTCCG AACGCAACTT
        80          90         100         110         120         130         140
CGCCCTGCTT GAGCGAGGCT GCGGTTTCCG AGGCCCTCTC CAGCCAAGGA AAAGCTACAC AAAAAGCCTG
       150         160         170         180         190         200         210
GATCACTCAT CGAACCACCC CTGAAGCCAG TGAAGGCTCT CTCGCCTCGC CCTCTAGCGT TCGTCTGGAG
       220         230         240         250         259         268
TAGCGCCACC CCGGCTTCCT GGGGACACAG GGTTGGCACC ATG GGG CCC ACC AGC GTC CCG
                                    ======  MET Gly Pro Thr Ser Val Pro

    277         286         295         304         313         322
CTG GTC AAG GCC CAC CGC AGC TCG GTC TCT GAC TAC GTC AAC TAT GAT ATC ATC
Leu Val Lys Ala His Arg Ser Ser Val Ser Asp Tyr Val Asn Tyr Asp Ile Ile

    331         340         349         358         367         376
GTC CGG CAT TAC AAC TAC ACG GGA AAG CTG AAT ATC AGC GCG GAC AAG GAG AAC
Val Arg His Tyr Asn Tyr Thr Gly Lys Leu Asn Ile Ser Ala Asp Lys Glu Asn
                 *                       *

    385         394         403         412         421         430
AGC ATT AAA CTG ACC TCG GTG GTG TTC ATT CTC ATC TGC TGC TTT ATC ATC CTG
Ser Ile Lys Leu Thr Ser Val Val Phe Ile Leu Ile Cys Cys Phe Ile Ile Leu

    439         448         457         466         475         484
GAG AAC ATC TTT GTC TTG CTG ACC ATT TGG AAA ACC AAG AAA TTC CAC CGA CCC
Glu Asn Ile Phe Val Leu Leu Thr Ile Trp Lys Thr Lys Lys Phe His Arg Pro

    493         502         511         520         529         538
ATG TAC TAT TTT ATT GGC AAT CTG GCC CTC TCA GAC CTG TTG GCA GGA GTA GCC
MET Tyr Tyr Phe Ile Gly Asn Leu Ala Leu Ser Asp Leu Leu Ala Gly Val Ala
```

FIG. 5A

```
    547         556         565         574         583         592
TAC ACA GCT AAC CTG CTC TTG TCT GGG GCC ACC ACC TAC AAG CTC ACT CCC GCC
Tyr Thr Ala Asn Leu Leu Leu Ser Gly Ala Thr Thr Tyr Lys Leu Thr Pro Ala

    601         610         619         628         637         646
CAG TGG TTT CTG CGG GAA GGG AGT ATG TTT GTG GCC CTG TCA GCC TCC GTG TTC
Gln Trp Phe Leu Arg Glu Gly Ser MET Phe Val Ala Leu Ser Ala Ser Val Phe

    655         664         673         682         691         700
AGT CTC CTC GCC ATC GCC ATT GAG CGC TAT ATC ACA ATG CTG AAA ATG AAA CTC
Ser Leu Leu Ala Ile Ala Ile Glu Arg Tyr Ile Thr MET Leu Lys MET Lys Leu

    709         718         727         736         745         754
CAC AAC GGG AGC AAT AAC TTC CGC CTC TTC CTG CTA ATC AGC GCC TGC TGG GTC
His Asn Gly Ser Asn Asn Phe Arg Leu Phe Leu Leu Ile Ser Ala Cys Trp Val

    763         772         781         790         799         808
ATC TCC CTC ATC CTG GGT GGC CTG CCT ATC ATG GGC TGG AAC TGC ATC AGT GCG
Ile Ser Leu Ile Leu Gly Gly Leu Pro Ile MET Gly Trp Asn Cys Ile Ser Ala

    817         826         835         844         853         862
CTG TCC AGC TGC TCC ACC GTG CTG CCG CTC TAC CAC AAG CAC TAT ATC CTC TTC
Leu Ser Ser Cys Ser Thr Val Leu Pro Leu Tyr His Lys His Tyr Ile Leu Phe

    871         880         889         898         907         916
TGC ACC ACG GTC TTC ACT CTG CTT CTG CTC TCC ATC GTC ATT CTG TAC TGC AGA
Cys Thr Thr Val Phe Thr Leu Leu Leu Leu Ser Ile Val Ile Leu Tyr Cys Arg
```

FIG. 5B

```
    925         934         943         952         961         970
ATC TAC TCC TTG GTC AGG ACT CGG AGC CGC CGC CTG ACG TTC CGC AAG AAC ATT
Ile Tyr Ser Leu Val Arg Thr Arg Ser Arg Arg Leu Thr Phe Arg Lys Asn Ile
                    O       O   •   O   O       •       O   O

    979         988         997        1006        1015        1024
TCC AAG GCC AGC CGC AGC TCT GAG AAT GTG GCG CTG CTC AAG ACC GTA ATT ATC
Ser Lys Ala Ser Arg Ser Ser Glu Asn Val Ala Leu Leu Lys Thr Val Ile Ile
     O           O                                       O

   1033        1042        1051        1060        1069        1078
GTC CTG AGC GTC TTC ATC GCC TGC TGG GCA CCG CTC TTC ATC CTG CTC CTG CTG
Val Leu Ser Val Phe Ile Ala Cys Trp Ala Pro Leu Phe Ile Leu Leu Leu Leu

   1087        1096        1105        1114        1123        1132
GAT GTG GGC TGC AAG GTG AAG ACC TGT GAC ATC CTC TTC AGA GCG GAG TAC TTC
Asp Val Gly Cys Lys Val Lys Thr Cys Asp Ile Leu Phe Arg Ala Glu Tyr Phe

   1141        1150        1159        1168        1177        1186
CTG GTG TTA GCT GTG CTC AAC TCC GGC ACC AAC CCC ATC ATT TAC ACT CTG ACC
Leu Val Leu Ala Val Leu Asn Ser Gly Thr Asn Pro Ile Ile Tyr Thr Leu Thr

   1195        1204        1213        1222        1231        1240
AAC AAG GAG ATG CGT CGG GCC TTC ATC CGG ATC ATG TCC TGC TGC AAG TGC CCG
Asn Lys Glu MET Arg Arg Ala Phe Ile Arg Ile MET Ser Cys Cys Lys Cys Pro

   1249        1258        1267        1276        1285        1294
AGC GGA GAC TCT GCT GGC AAA TTC AAG CGA CCC ATC ATC GCC GGC ATG GAA TTC
Ser Gly Asp Ser Ala Gly Lys Phe Lys Arg Pro Ile Ile Ala Gly MET Glu Phe
```

FIG. 5C

```
   1303          1312          1321          1330          1339          1348
AGC CGC AGC AAA TCG GAC AAT TCC TCC CAC CCC CAG AAA GAC GAA GGG GAC AAC
Ser Arg Ser Lys Ser Asp Asn Ser Ser His Pro Gln Lys Asp Glu Gly Asp Asn

   1357          1366          1375          1384          1393          1406
CCA GAG ACC ATT ATG TCT TCT GGA AAC GTC AAC TCT TCT TCC TAG AACTGGAAGC
Pro Glu Thr Ile MET Ser Ser Gly Asn Val Asn Ser Ser Ser

      1416          1426          1436          1446          1456          1466          1476
TGTCCACCCA CCGGAAGCGC TCTTTACTTG GTCGCTGGCC ACCCCAGTGT TTGGAAAAAA ATCTCTGGGC
      1486          1496          1506          1516          1526          1536          1546
TTCGACTGCT GCCAGGGAGG AGCTGCTGCA AGCCAGAGGG AGGAAGGGGG AGAATACGAA CAGCCTGGTG
      1556          1566          1576          1586          1596          1606          1616
GTGTCGGGTG TTGGTGGGTA GAGTTAGTTC CTGTGAACAA TGCACTGGGA AGGGTGGAGA TCAGGTCCCG
      1626          1636          1646          1656          1666          1676          1686
GCCTGGAATA TATATTCTAC CCCCCTGGAG CTTTGATTTT GCACTGAGCC AAAGGTCTAG CATTGTCAAG
      1696          1706          1716          1726          1736          1746          1756
CTCCTAAAGG GTTCATTTGG CCCCTCCTCA AAGACTAATG TCCCCATGTG AAAGCGTCTC TTTGTCTGGA
      1766          1776          1786          1796          1806          1816          1826
GCTTTGAGGA GATGTTTTCC TTCACTTTAG TTTCAAACCC AAGTGAGTGT GTGCACTTCT GCTTCTTTAG
      1836          1846          1856          1866          1876          1886          1896
GGATGCCCTG TACATCCCAC ACCCCACCCT CCCTTCCCTT CATACCCCTC CTCAACGTTC TTTTACTTTA
      1906          1916          1926          1936          1946          1956          1966
TACTTTAACT ACCTGAGAGT TATCAGAGCT GGGGTTGTGG AATGATCGAT CATCTATAGC AAATAGGCTA
      1976          1986          1996          2006          2016          2026          2036
TGTTGAGTAC GTAGGCTGTG GGAAGATGAA GATGGTTTGG AGGTGTAAAA CAATGTCCTT CGCTGAGGCC
```

FIG. 5D

```
      2046        2056        2066        2076        2086        2096        2106
AAAGTTTCCA TGTAAGCGGG ATCCGTTTTT TGGAATTTGG TTGAAGTCAC TTTGATTTCT TTAAAAAACA
      2116        2126        2136        2146        2156        2166        2176
TCTTTTCAAT GAAATGTGTT ACCATTTCAT ATCCATTGAA GCCGAAATCT GCATAAGGAA GCCCACTTTA
      2186        2196        2206        2216        2226        2236        2246
TCTAAATGAT ATTAGCCAGG ATCCTTGGTG TCCTAGGAGA AACAGACAAG CAAAACAAAG TGAAAACCGA
      2256        2266        2276        2286        2296        2306        2316
ATGGATTAAC TTTTGCAAAC CAAGGGAGAT TTCTTAGCAA ATGAGTCTAA CAAATATGAC ATCCGTCTTT
      2326        2336        2346        2356        2366        2376        2386
CCCACTTTTG TTGATGTTTA TTTCAGAATC TTGTGTGATT CATTTCAAGC AACAACATGT TGTATTTTGT
      2396        2406        2416        2426        2436        2446        2456
TGTGTTAAAA GTACTTTTCT TGATTTTTGA ATGTATTTGT TTCAGGAAGA AGTCATTTTA TGGATTTTTC
      2466        2476        2486        2496        2506        2516        2526
TAACCCGTGT TAACTTTTCT AGAATCCACC CTCTTGTGCC CTTAAGCATT ACTTTAACTG GTAGGGAACG
      2536        2546        2556        2566        2576        2586        2596
CCAGAACTTT TAAGTCCAGC TATTCATTAG ATAGTAATTG AAGATATGTA TAAATATTAC AAAGAATAAA
      2606        2616        2626        2636        2646        2656        2666
AATATATTAC TGTCTCTTTA GTATGGTTTT CAGTGCAATT AAACCGAGAG ATGTCTTGTT TTTTTAAAAA
      2676        2686        2696        2706        2716        2726        2736
GAATAGTATT TAATAGGTTT CTGACTTTTG TGGATCATTT TGCACATAGC TTTATCAACT TTTAAACATT
      2746        2756        2766
AATAAACTGA TTTTTTTAAA GAAAAAAAAA AAAAAAAG
```

FIG. 5E

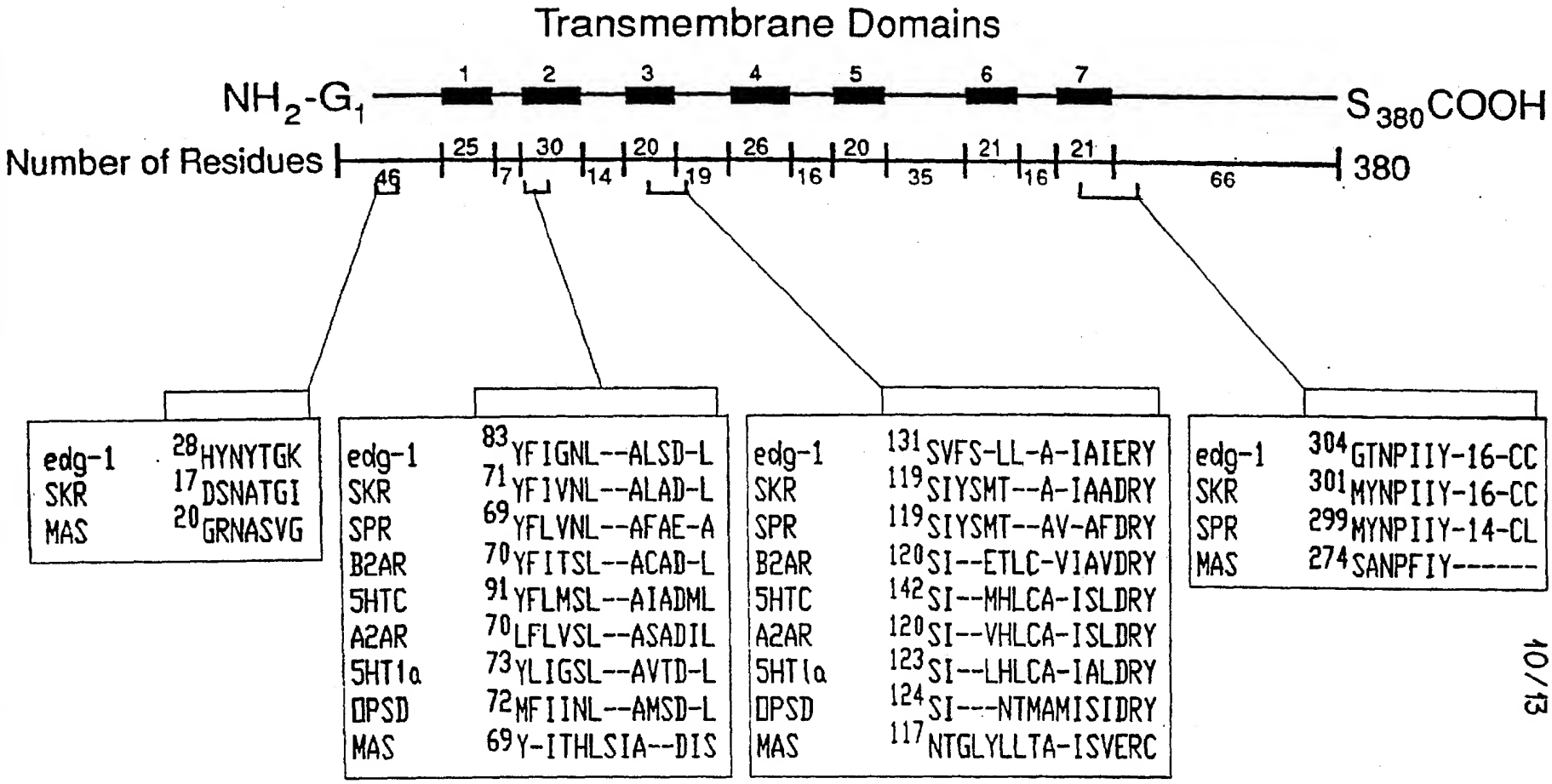

Transmembrane Domains
NH2-G1
1
2
3
4
5
6
7
S380COOH
Number of Residues
46
25
7
30
14
20
19
26
16
20
35
21
16
21
66
380
edg-1 28HYNYTGK
SKR 17DSNATGI
MAS 20GRNASVG
edg-1 83YFIGNL--ALSD-L
SKR 71YFIVNL--ALAD-L
SPR 69YFLVNL--AFAE-A
B2AR 70YFITSL--ACAD-L
5HTC 91YFLMSL--AIADML
A2AR 70LFLVSL--ASADIL
5HT1a 73YLIGSL--AVTD-L
PSD 72MFIINL--AMSD-L
MAS 69Y-ITHLSIA--DIS
edg-1 131SVFS-LL-A-IAIERY
SKR 119SIYSMT--A-IAADRY
SPR 119SIYSMT--AV-AFDRY
B2AR 120SI--ETLC-VIAVDRY
5HTC 142SI--MHLCA-ISLDRY
A2AR 120SI--VHLCA-ISLDRY
5HT1a 123SI--LHLCA-IALDRY
PSD 124SI---NTMAMISIDRY
MAS 117NTGLYLLTA-ISVERC
edg-1 304GTNPIIY-16-CC
SKR 301MYNPIIY-16-CC
SPR 299MYNPIIY-14-CL
MAS 274SANPFIY------

FIG. 6

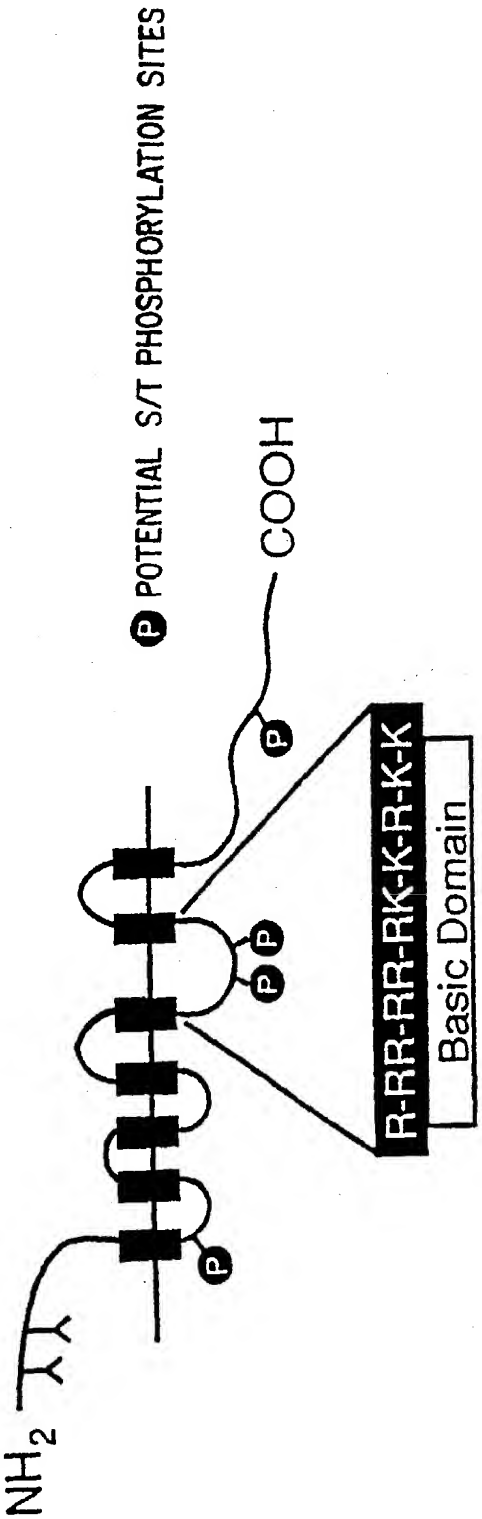

POTENTIAL S/T PHOSPHORYLATION SITES
COOH
NH2
R-RR-RR-RK-K-R-K-K
Basic Domain

FIG. 7

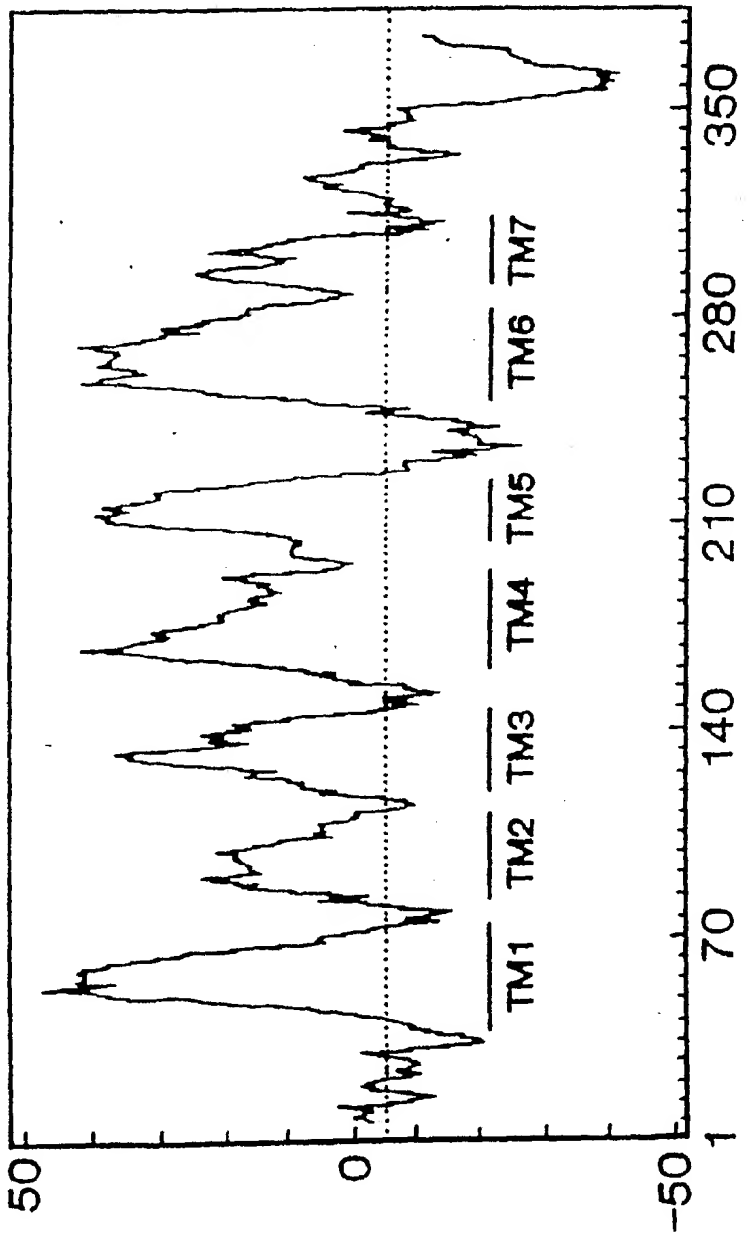

50
0
-50
1
70
140
210
280
350
Residue
TM1
TM2
TM3
TM4
TM5
TM6
TM7
Hydrophobicity

FIG. 8

FIG. 9B

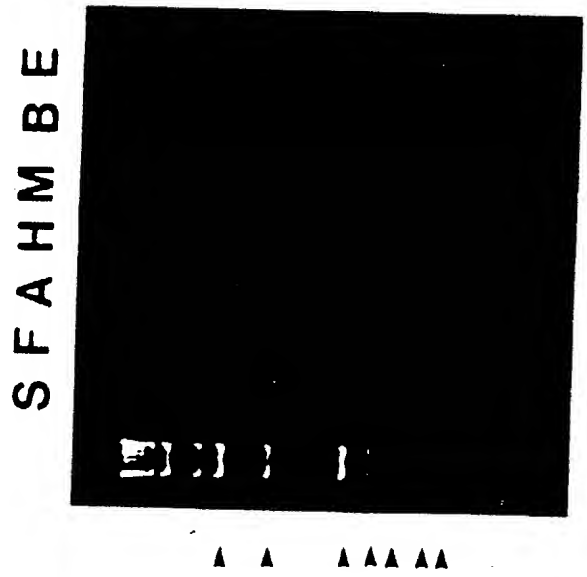
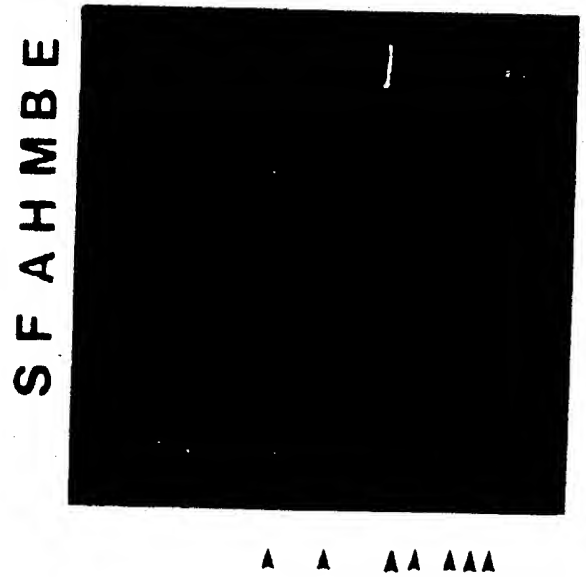


FIG. 9A



INTERNATIONAL SEARCH REPORT

International Application No PCT/US91/02344

I. CLASSIFICATION OF SUBJECT MATTER (if several classification symbols apply, indicate all) [5]

According to International Patent Classification (IPC) or to both National Classification and IPC

IPC(5): C12N 15/12; C07K 15/06, 15/14
U.S.CL.: 536/27, 530/350,395

II. FIELDS SEARCHED

Minimum Documentation Searched [7]

Classification System	Classification Symbols
U.S.Cl.	536/27; 530/350,395

Documentation Searched other than Minimum Documentation to the Extent that such Documents are Included in the Fields Searched [8]

APS AND DIALOG Files 357,155,WPI,72,35,5 and 399 searched for edg type receptor proteins and sequences.

III. DOCUMENTS CONSIDERED TO BE RELEVANT [9]

Category *	Citation of Document, [11] with indication, where appropriate, of the relevant passages [12]	Relevant to Claim No. [13]
X,P	Journal of Biological Chemistry, vol. 265, No. 16, issued 05 June 1990. Hla et al., "An abundant transcript induced in differentiating human endothelial cells encodes a polypeptide with structural similarities to G-protein-coupled receptors", pages 9308-9313. See whole publication, especially the abstract, p. 9309 and 9311.	1-8
X	Science, vol. 241, issued 16 September 1988. Klein et al., "A chemoattractant receptor controls development in Dictyostelium discoideum", pages 1467-1472. See whole publications, especially Figure 8 on p. 1472.	3-8
A	Science, vol. 245, issued 08 September 1989, Devreotes. "Dictyostelium discoideum: a model system for cell-cell interactions in development". pages 1054-1058. See whole publication.	1-8

* Special categories of cited documents: [10]

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

IV. CERTIFICATION

Date of the Actual Completion of the International Search	Date of Mailing of this International Search Report
02 July 1991	26 JUL 1991

International Searching Authority	Signature of Authorized Officer
ISA/US	Keith C. Furman KEITH C. FURMAN

Form PCT/ISA/210 (second sheet)